Integrable equations as compatibility conditions.

The Lax formalism for the KdV equation. Consider the differential operators

$$L = -6\frac{d^2}{dx^2} - u\,, \qquad B = -4\frac{d^3}{dx^3} - u\frac{d}{dx} - \frac{1}{2}u_x\,.$$

L is symmetric, and B is skew-symmetric. Then it was noted by Peter Lax in 1968 that the operator equation

$$\frac{dL}{dt} + [L, B] = 0$$

is equivalent to the KdV equation $u_t + uu_x + u_{xxx} = 0$ in the sense that both sides of the equation turn out to be operators of multiplication by a function. Here $[L, B] = LB - BL$ is the operator commutator. This is a direct calculation. An operator equation of this form is today called a *Lax equation.*

As we will soon see in a concrete example, this form of the KdV equation *immediately* shows (among other things) that the eigenvalues of L are independent of t. Another way to derive this form of the KdV equation is as the *compatibility condition* of two linear problems where now we *assume* that λ is a fixed parameter:

$$L\phi = \lambda\phi\,, \qquad \text{and} \qquad \phi_t = B\phi\,.$$

Indeed, from these two we have

$$\frac{\partial}{\partial t}(L\phi) = \frac{dL}{dt}\phi + L\phi_t = \frac{dL}{dt}\phi + LB\phi\,,$$

and also

$$\frac{\partial}{\partial t}(L\phi) = \frac{\partial}{\partial t}(\lambda\phi) = \lambda\phi_t = \lambda B\phi = B(\lambda\phi) = BL\phi\,.$$

Therefore as ϕ is general, we obtain

$$\frac{dL}{dt} + LB = BL\,.$$

The key importance of Lax's observation is that *any* equation that can be cast into such a framework for other operators L and B has automatically many of the features of the KdV equation, including an infinite number of local conservation laws.

Rewriting KdV as a zero-curvature condition. Going further with the point of view of looking at integrable nonlinear problems as the compatibility conditions of two linear problems (frequently involving an arbitrary "spectral" parameter λ), we can go from linear operators to matrices at the cost of introducing some powers of λ. To do this, note that as the eigenvalue equation is second order, we may easily write it in first-order form by introducing

$$\phi_1 = \phi\,, \qquad \text{and} \qquad \phi_2 = \phi_x\,,$$

So $L\phi = \lambda\phi$ can be rewritten as

$$\frac{\partial}{\partial x}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} = \begin{bmatrix}0 & 1\\ -(u+\lambda)/6 & 0\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} = \mathbf{U}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix}.$$

Similarly, we have

$$\begin{aligned}
\phi_{1t} &= \phi_t\\
&= -4\phi_{xxx} - u\phi_x - \frac{1}{2}u_x\phi\\
&= -4(-(u+\lambda)\phi/6)_x - u\phi_x - \frac{1}{2}u_x\phi\\
&= \frac{1}{6}u_x\phi + \left(\frac{2}{3}\lambda - \frac{1}{3}u\right)\phi_x\\
&= \frac{1}{6}u_x\phi_1 + \left(\frac{2}{3}\lambda - \frac{1}{3}u\right)\phi_2\,.
\end{aligned}$$

And therefore

$$\begin{aligned}
\phi_{2t} &= \phi_{1tx} \\
&= \frac{1}{6}u_x\phi_{1x} + \frac{1}{6}u_{xx}\phi_1 + \left(\frac{2}{3}\lambda - \frac{1}{3}u\right)\phi_{2x} - \frac{1}{3}u_x\phi_2 \\
&= \frac{1}{6}u_x\phi_2 + \frac{1}{6}u_{xx}\phi_1 - \left(\frac{2}{3}\lambda - \frac{1}{3}u\right)\left(\frac{1}{6}\lambda + \frac{1}{6}u\right)\phi_1 - \frac{1}{3}u_x\phi_2 \\
&= \left(-\frac{1}{9}\lambda^2 - \frac{1}{18}\lambda u + \frac{1}{18}u^2 + \frac{1}{6}u_{xx}\right)\phi_1 - \frac{1}{6}u_x\phi_2\,.
\end{aligned}$$

Or, writing as a first-order system,

$$\frac{\partial}{\partial t}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} = \begin{bmatrix}\frac{1}{6}u_x & \frac{2}{3}\lambda - \frac{1}{3}u \\ -\frac{1}{9}\lambda^2 - \frac{1}{18}\lambda u + \frac{1}{18}u^2 + \frac{1}{6}u_{xx} & -\frac{1}{6}u_x\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} = \mathbf{V}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix}.$$

The compatibility condition now takes the form of a matrix equation, the *zero-curvature condition*:

$$\frac{\partial \mathbf{U}}{\partial t} - \frac{\partial \mathbf{V}}{\partial x} + [\mathbf{U}, \mathbf{V}] = \mathbf{0}\,.$$

If we separate out the coefficients of powers of λ in this equation then the coefficients of λ^3, λ^2, and λ all vanish identically. The constant term is the matrix equation

$$\begin{bmatrix}0 & 0\\ -\frac{1}{6}(u_t + uu_x + u_{xxx}) & 0\end{bmatrix} = \mathbf{0}\,.$$

The name comes from the following geometrical interpretation. The equations $(\partial_x - \mathbf{U})\phi = 0$ and $(\partial_t - \mathbf{V})\phi = 0$ define a *connection* on a two-dimensional *vector bundle* over the (x, t)-plane. The first equation describes how to "parallel-translate" a vector ϕ in the x-direction, and the second equation describes how to "parallel-translate" a vector ϕ in the t-direction. The matrices $\mathbf{U}$ and $\mathbf{V}$ are the *connection coefficients.* A connection is said to have *zero curvature* if parallel translation of a vector ϕ along a path from a point (x_0, t_0) to another point (x_1, t_1) gives the same result independent of path connecting the points. This is the same thing as asserting the existence of a full two-dimensional basis of simultaneous solutions of the equations $(\partial_x - \mathbf{U})\phi = 0$ and $(\partial_t - \mathbf{V})\phi = 0$, which is the above zero-curvature condition that must be satisfied by the connection coefficients. Therefore, every solution of the KdV equation defines a connection with zero curvature.

Generalization. The two linear equations for ϕ making up a zero-curvature connection are said to form a *Lax pair.* This idea is one of the most central ones in the theory of integrable systems: each integrable nonlinear problem can be represented as the compatibility condition between two linear equations of a Lax pair. Here are some other examples.

A common choice for the matrix $\mathbf{U}$ is the "AKNS" (Ablowitz, Kaup, Newell, and Segur) spectral problem:

$$\mathbf{U} = \begin{bmatrix}i\lambda & q\\ r & -i\lambda\end{bmatrix} = \lambda\mathbf{U}_1 + \mathbf{U}_0\,.$$

Here q and r are some functions of x and t. We will derive nonlinear equations for them by seeking matrices $\mathbf{V}$ giving rise to a zero-curvature connection. For example, suppose we seek $\mathbf{V}$ as a quadratic polynomial in λ:

$$\mathbf{V} = \lambda^2\mathbf{V}_2 + \lambda\mathbf{V}_1 + \mathbf{V}_0\,.$$

Then, the zero-curvature condition splits into several equations corresponding to powers of λ:

$$\begin{aligned}
\text{coefficient of } \lambda^3\text{:}\quad & [\mathbf{U}_1, \mathbf{V}_2] = \mathbf{0}\,, \\
\text{coefficient of } \lambda^2\text{:}\quad & -\frac{\partial \mathbf{V}_2}{\partial x} + [\mathbf{U}_1, \mathbf{V}_1] + [\mathbf{U}_0, \mathbf{V}_2] = \mathbf{0}\,, \\
\text{coefficient of } \lambda\text{:}\quad & \frac{\partial \mathbf{U}_1}{\partial t} - \frac{\partial \mathbf{V}_1}{\partial x} + [\mathbf{U}_1, \mathbf{V}_0] + [\mathbf{U}_0, \mathbf{V}_1] = \mathbf{0}\,, \\
\text{constant term:}\quad & \frac{\partial \mathbf{U}_0}{\partial t} - \frac{\partial \mathbf{V}_0}{\partial x} + [\mathbf{U}_0, \mathbf{V}_0] = \mathbf{0}\,.
\end{aligned}$$

Note that $\mathbf{U}_1 = i\sigma_3$, where the Pauli matrix σ_3 is defined by

$$\sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

We now consider these four equations one at a time. The λ^3 equation can be solved by choosing $\mathbf{V}_2 = \mathbf{U}_1 = i\sigma_3$ (more generally an arbitrary diagonal matrix would do).

With this choice the λ^2 equation becomes

$$[i\sigma_3, \mathbf{V}_1 - \mathbf{U}_0] = 0\,.$$

Note that for a general matrix $\mathbf{A}$, we have $[i\sigma_3, \mathbf{A}] = 2i\sigma_3 \mathbf{A}^{OD}$, where $\mathbf{A}^{OD}$ is the off-diagonal part of $\mathbf{A}$. Therefore since σ_3 is invertible, this says that $\mathbf{V}_1$ must differ from $\mathbf{U}_0$ by a diagonal matrix. But let's solve this by also choosing that $\mathbf{V}_1 = \mathbf{U}_0$.

The λ equation then becomes

$$-\frac{\partial \mathbf{U}_0}{\partial x} + [i\sigma_3, \mathbf{V}_0] = \mathbf{0}\,.$$

This equation determines the off-diagonal part of $\mathbf{V}_0$:

$$\mathbf{V}_0^{OD} = -\frac{i}{2}\sigma_3 \frac{\partial \mathbf{U}_0}{\partial x} = \begin{bmatrix} 0 & -iq_x/2 \\ ir_x/2 & 0 \end{bmatrix}.$$

Finally, we consider the constant term equation which we split into diagonal and off-diagonal parts: the diagonal part is

$$-\frac{\partial \mathbf{V}_0^D}{\partial x} + [\mathbf{U}_0, \mathbf{V}_0^{OD}] = \mathbf{0}$$

and the off-diagonal part is

$$\frac{\partial \mathbf{U}_0}{\partial t} - \frac{\partial \mathbf{V}_0^{OD}}{\partial x} + [\mathbf{U}_0, \mathbf{V}_0^D] = \mathbf{0}\,.$$

Inserting our formula for $\mathbf{V}_0^{OD}$ into the diagonal part gives

$$\frac{\partial \mathbf{V}_0^D}{\partial x} = \begin{bmatrix} iqr_x/2 + irq_x/2 & 0 \\ 0 & -irq_x/2 - iqr_x/2 \end{bmatrix}$$

so we may solve this by choosing

$$\mathbf{V}_0^D = \begin{bmatrix} iqr/2 & 0 \\ 0 & -iqr/2 \end{bmatrix}.$$

The matrices $\mathbf{V}_2$, $\mathbf{V}_1$, and $\mathbf{V}_0$ have thus been determined and the only part of the zero-curvature condition that remains is the off-diagonal part of the constant term equation.

The off-diagonal part of the constant term equation reads as follows:

$$\begin{bmatrix} 0 & q_t + iq_{xx}/2 - iq^2 r \\ r_t - ir_{xx}/2 + ir^2 q & 0 \end{bmatrix} = \mathbf{0}\,.$$

In other words, the connection defined by the coefficient matrices $\lambda\mathbf{U}_1 + \mathbf{U}_0$ and $\lambda^2\mathbf{V}_2 + \lambda\mathbf{V}_1 + \mathbf{V}_0$ will have zero curvature if and only if the functions q and r satisfy the nonlinear equations

$$iq_t - \frac{1}{2}q_{xx} + q^2 r = 0\,, \qquad \text{and} \qquad -ir_t - \frac{1}{2}r_{xx} + r^2 q = 0\,.$$

Generally this is a coupled system. However, it is consistent with the relation $r = \pm q^*$, which gives

$$iq_t - \frac{1}{2}q_{xx} \pm |q|^2 q = 0\,, \qquad \text{and} \qquad -iq_t^* - \frac{1}{2}q_{xx}^* \pm |q|^2 q^*\,,$$

so the second equation is just the complex conjugate of the first. Therefore we have found the Lax pair representation for the focusing and defocusing nonlinear Schrödinger equations.

As a second example using the same AKNS spectral problem, seek $\mathbf{V}$ in the form:

$$\mathbf{V} = \lambda^{-1}\mathbf{V}_{-1}\,.$$

The zero-curvature equation then splits into three parts:

$$\text{coefficient of } \lambda\text{:} \qquad \frac{\partial \mathbf{U}_1}{\partial t} = \mathbf{0}\,,$$

which is trivially satisfied since $\mathbf{U}_1$ is constant,

$$\text{constant terms:} \qquad \frac{\partial \mathbf{U}_0}{\partial t} + [\mathbf{U}_1, \mathbf{V}_{-1}] = \mathbf{0}\,,$$

and

$$\text{coefficient of } \lambda^{-1}\text{:} \qquad -\frac{\partial \mathbf{V}_{-1}}{\partial x} + [\mathbf{U}_0, \mathbf{V}_{-1}] = \mathbf{0}\,.$$

The constant terms determine the off-diagonal part of $\mathbf{V}_{-1}$:

$$2i\sigma_3 \mathbf{V}_{-1}^{OD} = -\frac{\partial \mathbf{U}_0}{\partial t}\,, \qquad \text{so} \qquad \mathbf{V}_{-1}^{OD} = \frac{i}{2}\sigma_3 \frac{\partial \mathbf{U}_0}{\partial t} = \begin{bmatrix} 0 & iq_t/2 \\ -ir_t/2 & 0 \end{bmatrix}.$$

The diagonal part of the λ^{-1} equation is

$$-\frac{\partial \mathbf{V}_{-1}^{D}}{\partial x} + \begin{bmatrix} -i(qr)_t/2 & 0 \\ 0 & i(qr)_t \end{bmatrix} = \mathbf{0}\,,$$

which says in particular that $\mathbf{V}_{-1}^{D} = \alpha\sigma_3 + c\mathbb{I}$, where c has to be a constant but α is a function of x and t related to q and r by

$$\alpha_x + \frac{1}{2} i(qr)_t = 0\,.$$

The off-diagonal part of the λ^{-1} equation is

$$-\begin{bmatrix} 0 & iq_{xt}/2 \\ -ir_{xt}/2 & \end{bmatrix} + [\mathbf{U}_0, \mathbf{V}_{-1}^{D}] = \mathbf{0}\,.$$

This tells us that

$$\frac{1}{2} iq_{xt} - 2\alpha q = 0\,, \qquad \text{and} \qquad -\frac{1}{2} ir_{xt} + 2\alpha r = 0\,.$$

So the zero curvature condition amounts to three equations on three unknowns. A special case is obtained by setting

$$\alpha = -\frac{1}{4} i\cos(u)\,, \qquad \text{and} \qquad q = -r = \frac{1}{2} u_x\,,$$

upon which we find that either $u_x = 0$ or

$$u_{xt} = \sin(u)\,.$$

This is a form of the sine-Gordon equation. Indeed, if we introduce new coordinates by $\xi = x + t$ and $\tau = x - t$, then

$$u_{xt} = \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \tau}\right)\left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \tau}\right) u = u_{\xi\xi} - u_{\tau\tau}\,.$$